UNIVERSIDAD NACIONAL DE COLOMBIA

Machine Learning Lab 2.2.

Miguel Andrés Villamil Carrillo

Universidad Nacional de Colombia
Bogotá, Colombia
2018

1 Original algorithm

Algorithm 1 Apply MultinomialNB

procedure ApplyMultinomialNB($\mathbb{C}, V, prior, condprob, d$)
  $W \leftarrow ExtractTokensFromDoc(V, d)$
  for each $c \in \mathbb{C}$
  do $score[c] \leftarrow log\ prior[c]$
    for each $t \in W$
    do $score[c]$ += $log\ condprob[t][c]$
  return $argmax_{c \in \mathbb{C}}\ score[c]$

2 Proposed algorithm

Algorithm 2 Apply MultinomialNB

$WDistinct \leftarrow ExtractDistinctTokensFromDoc(V, d)$
$W \leftarrow ExtractTokensFromDoc(V, d)$
for each $c \in \mathbb{C}$
do $score[c] \leftarrow log\ prior[c]$
  for each $t \in WDistinct$
  do $t_c \leftarrow CountTokensOfTerm(t, W)$
  do $score[c]$ += $log\ condprob[t][c] * t_c$
return $argmax_{c \in \mathbb{C}}\ score[c]$